Biomedical Signal Processing and Control 18 (2015) 69–79

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Automatic sleep staging by simultaneous analysis of ECG and respiratory signals in long epochs

Farideh Ebrahimi [a,*], Seyed-Kamaledin Setarehdan [a], Homer Nazeran [b]

[a] Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran

[b] Department of Electrical and Computer Engineering, The University of Texas at El Paso, El Paso, TX, USA

ARTICLE INFO

Article history:
Received 29 July 2014
Received in revised form
25 November 2014
Accepted 8 December 2014

Keywords:
Heart Rate Variability (HRV)
ECG-Derived Respiration (EDR)
Empirical Mode Decomposition (EMD)
Discrete Wavelet Transform (DWT)
Detrended Fluctuation Analysis (DFA)
Support Vector Machine-Recursive Feature Elimination (SVM-RFE)

ABSTRACT

EEG, EMG, and EOG are very informative signals recorded in polysomnography (PSG) and used for sleep staging. Their reliable acquisition at home, however, is difficult. In comparison, ECG and thoracic respiratory (R) signals are easier to record and can be useful in home sleep monitoring systems. The simultaneous utilization of Heart Rate Variability (HRV) and respiratory (R) signals seems a plausible scenario as both heart rate (HR) and respiration rate (RR) vary during different sleep states. Therefore, we explored the combined discriminative capacity (accuracy, sensitivity, and specificity) of ECG/R signals in automatic sleep staging. As baseline, we classified the wakefulness, Stage 2, SWS (slow wave sleep) and REM sleep by using a Support Vector Machine (SVM) fed with a set of features extracted from: (a) HRV (34-features), (b) HRV/ECG-Derived Respiration (45-features), and (c) combined HRV/R (45-features) signals. Approach (a) produced reasonable discriminative capacity, while approach (b) significantly improved the classification; however, the best outcomes were achieved by using approach (c). Then, we enhanced the SVM classifier with the Recursive Feature Elimination (RFE) method. The classification results were improved with 35 out of the 45 HRV/RS-EDR features. In comparison, best results were obtained by combining 27 out of the 45 features derived from HRV/R signals, in which the optimal feature set selected by the SVM-RFE method, included a combination of time domain, time-frequency, and fractal features, as well as entropies. Overall, these improvements revealed that it is possible to simplify home monitoring of sleep disorders and achieve high discriminative capacity (accuracy = 89.32%, specificity = 92.88%, and sensitivity = 78.64%) in automatic sleep staging by the exclusive recording of cardiorespiratory signals.

1. Introduction

Generally, there are three different states of existence: wakefulness, rapid eye movement (REM) sleep and non-rapid eye movement (NREM) sleep. According to the latest standard of the American Academy of Sleep Medicine (AASM), NREM sleep states are divided into three stages (stages 3 and 4 in the Rechtschaffen and Kales [1] standard are scored as SWS, in the AASM standard) [2]. EEG (Electroencephalogram), EMG (Electromyogram) and EOG (Electro-oculogram) signals are among the most important polysomnography (PSG) signals used for sleep staging. High performance can be reached with such recordings but the reliable acquisition of these signals is not easily possible in the home environment. Instead, ECG (Electrocardiogram) and respiration signals, which are much easier to record, can be useful in home sleep monitoring systems.

It has been shown that autonomic nervous system (ANS) activity differs during wakeful and sleep states. During NREM sleep, sympathetic input is reduced and parasympathetic activity predominates, resulting in relative stability in the ANS function. REM sleep, in contrast, is characterized by a highly variable sympathetic activation (which can be comparable to or higher than wakefulness levels) punctuated by a phasic parasympathetic discharge, which results in fluctuating cardiovascular and respiratory behaviors. Accelerations and pauses in heart rate and irregular ventilation are therefore characteristic of REM sleep [3]. In addition, by directly recording sympathetic nerve activity for several hours, investigators have provided solid evidence that sympathetic activity is reduced by more than half from wakefulness to SWS stage but increased to levels above waking values during REM sleep [4].

* Corresponding author at: School of Electrical and Computer Engineering, University College of Engineering, University of Tehran, North Kargar St., Tehran 1943883654, Iran. Tel.: +98 21 6111 4177; fax: +98 21 8863 3029.
E-mail address: f.ebrahimi88@ut.ac.ir (F. Ebrahimi).

http://dx.doi.org/10.1016/j.bspc.2014.12.003
1746-8094/

Throughout NREM sleep, respiration is controlled entirely automatically, functioning primarily to maintain the level of carbon dioxide in the blood at a slightly higher level than the awake stage and, to a lesser extent, maintain oxygen in the blood at a slightly lower level. There is about a 15% decrease in air volume entering and leaving the lungs per minute during NREM sleep. Overall, respiration, especially during SWS stage, is regular and slightly deeper. The control of respiration during REM sleep is another story. The respiration rate and depth are very irregular and a rapid and shallow pattern tends to prevail. The average air intake and the level of oxygen in the blood can be about the same or less than that of NREM sleep. There may even be pauses in respiration during REM sleep [5].

The sleep literature shows that the identification of sleep stages from the ECG is the most recent endeavor in the field and numerous papers have shown that sleep staging depends on ECG-related parameters. For example, Vigo et al. showed the importance of feature extraction from 5-min epochs of HRV (Heart Rate Variability) signals in sleep staging [6]. In addition, it has been shown that the ratio of the power in the LF (low frequency) and HF (high frequency) bands of the heart rate is a valuable feature in sleep staging [7]. Moreover, Penzel et al., by using Detrended Fluctuation Analysis (DFA), showed that the dynamics of HRV signals are different in the wakeful state and sleep stages [8]. They also used spectral analysis and DFA in order to extract information from HRV signals separately for automatic sleep staging and showed that DFA is a better approach than spectral analysis [9]. The importance of R signals has also been probed in some studies. Carskadon et al., for instance, evaluated the use of R signal variability in wakeful and sleep stages in children [10]. They observed that the respiration rate and its regularity decreased from wakefulness to NREM sleep. Miyata et al. showed that R signals were generated from a nonlinear underlying system [11]. Another study evaluated the importance of using approximate entropy of R signals in wakeful and sleep stages [12]. Generally speaking, having nonlinear characteristics does not necessarily mean that the signal has fractal characteristics, however some studies have shown that R signals could exhibit fractal behavior [13–15]. Ordinarily, cardiac and respiratory rhythms can synchronize in different ratios. For some synchronization ratios, a pronounced sleep-stage dependency has been observed with the low degree of synchronization during REM and Wake, higher during LS??, and most pronounced during deep sleep [16,17]. Recent findings in this area have motivated some scholars to classify sleep stages automatically based on HRV [18–20], and by the combination of ECG and respiratory signals [21–23].

The objective of this paper is to report our ongoing research efforts on ECG-based automatic sleep staging. Here we evaluate the potential utility of combining ECG and R signals in automatic sleep staging. First, ECG-Derived Respiration (EDR) signals based on the R and S waves of the QRS complex, called the RS-EDR signal, as well as HRV signals, were derived from raw ECG data. Then, we extracted features from RS-EDR, raw R and HRV signals. Following this step, the variability of some of these features in different sleep and wakeful stages was compared. Finally, we performed automatic sleep staging by using HRV signal features alone, in combination with RS-EDR signal features, and then by combining R signal features with HRV signal features. A Support Vector Machine (SVM) classifier was implemented for classification and then a wrapper feature selection method, named SVM Recursive Feature Elimination (SVM-RFE), was used for feature optimization, before classification.

2. Materials and methods

Fig. 1 shows the block diagram for the automatic sleep staging algorithm used in our research. It comprises of four steps: (1) preprocessing, (2) feature extraction, (3) SVM-RFE-based feature reduction, and (4) SVM-based automatic sleep classification. Below, each block is described in more detail.

2.1. Polysomnographic data

The Sleep Heart Health Study (SHHS) database was used to provide the sleep data for this investigation [24]. These recordings were selected by considering a Respiratory Disturbance Index 3 Percent (RDI3P[1]) <5 to have near-normal characteristics. Subjects did not use beta-blockers, alpha-blockers or inhibitors. ECG signals from 30 overnight (nocturnal) polysomnographic recordings sampled at 250 Hz (Fs = sampling rate) and thoracic excursion data recorded by inductive plethysmography bands and sampled at 10 Hz were used as reference R signals. Sleep architecture for these data was determined in each subject according to Rechtschaffen and Kales (R&K) criteria in 0.5-min epochs [1]. For each recording, continuous segments of ECG signals in each sleep cycle were separated first and then, 5-min segments of different sleep stages were selected manually. In this study, a small part of wake stage, which happens during sleep times (or relaxed wakefulness) was considered. There were no continuous 5-min epochs available for Stage 1 in the database that we used. Since, the study was carried out based on using 5-min segments of cardiorespiratory signals; we did not consider Sleep Stage 1 for further analysis. In addition, we did not use some of the 5-min epochs for which the R-detection algorithm was not effective. The remaining 5-min epochs were selected randomly so that the number of feature vectors in the 4 classes would not be too different. Table 1 shows the number of 5-min segments of wakeful and different sleep stages for all recordings.

Moreover, atrial fibrillation frequently happens in patients with sleep apnea [25–30]. As in our study, the recordings were selected by considering a Respiratory Disturbance Index 3 Percent (RDI3P[1 1]) <5 to have near-normal characteristics, we assumed there were no atrial fibrillations in our database.

2.2. HRV signal derivation

Continuous 5-min segments of ECG signals were filtered by using an FIR band pass filter with low and high cut-off frequencies of 8 and 20 Hz, respectively [31]. In order to extract HRV signals, QRS complexes were detected first by using an Enhanced Hilbert Transform (EHT) algorithm [32] and then were quality-assured manually for missing beats correction. RRI (ECG R–R interval) time series were constructed from QRS complexes by using a cubic spline interpolation method with a 2.4 Hz sampling rate. As non-stationary characteristics in the form of *linear trends* or more complex *trends* disturb HRV signal analysis and cause unwanted effects on the extracted features from these signals, existing trends were removed from interpolated RRI time series [33].

2.3. ECG-Derived Respiration (EDR) signal

During breathing when the lungs fill with and empty of air, the ECG signal recorded from the surface of body is affected by the motion of electrodes with respect to the position of the heart and the variations in thoracic impedance. These effects often appear as ECG amplitude modulation, which is indicative of the R signal. The research literature abounds with efforts for extraction of the R signal from the ECG signal [34–39]. RS_{ampl} is a method of EDR extraction, which is presented by Mason et al. [35]. In this method,

[1] RDI3P was defined as the number of apneas and hypopneas per hour of sleep at 3% oxygen desaturation.

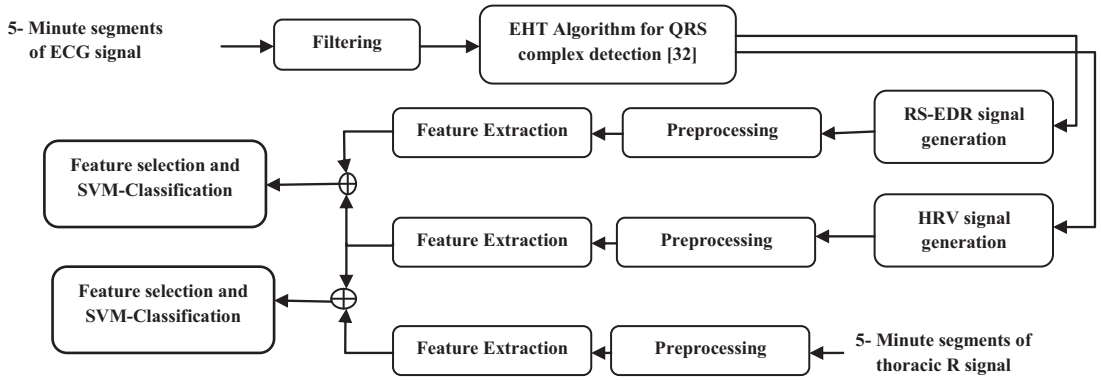

Fig. 1. Automatic sleep staging algorithm using ECG and thoracic R signals. EHT, RS-EDR and R, are Enhanced Hilbert Transform, ECG-Derived Respiration (EDR) signals based on the R and S waves of the QRS complex and respiratory signal, respectively.

the amplitude difference between the R wave and the S wave is calculated.

$$\text{ampl}(i) = r_{\text{ampl}}(i) - s_{\text{ampl}}(i) \quad (1)$$

In the above equation, r_{ampl} and s_{ampl} are R wave and S wave amplitudes, respectively. The s_{ampl} is the minimum amplitude of ECG signals in the 0.1 s window immediately following the R wave. The RS_{ampl} EDR (RS-EDR) method has been used in other research efforts for detection of sleep disorder breathing (SDB), as it is not sensitive to ECG baseline wander [37].

In our investigation here, the EDR signal was extracted by using the RS_{ampl} or RS-EDR method from ECG signals. The QRS complexes were detected first from 5-min segments of ECG signals in different sleep stages and wakefulness (as described in Section 2.2) by using an Enhanced Hilbert Transform (EHT) algorithm [32]. Then, each sample point of the RS-EDR signal was obtained by calculating the amplitude difference between the R and S waves (as explained above) and then saved as a vector. The RS-EDR signal is an unevenly sampled sequence, with samples corresponding to the QRS detection times. To generate a smooth EDR waveform, a 2-point moving average was applied to the extracted vector and the RS-EDR time series were interpolated using cubic splines to derive a signal uniformly sampled at 10 Hz. The resulting signal was then filtered within respiration rates of 5 and 25 cycles per minute using a Chebyshev Type I band pass filter. Fig. 2 shows the results of extraction of the R signal from a 5-min ECG signal by using the RS-EDR method (after IIR filtering).

2.4. Feature extraction from thoracic R and RS-EDR signals

2.4.1. Peak and valley detection in R signals (thoracic R and RS-EDR signals)

The peaks and valleys in R signals indicate the beginning of an expiration and an inspiration, respectively. However, the detection of peaks and valleys also includes the detection of local optima (upper trace of Fig. 3). To remove false inspirations and expirations from local optima and to separate global optima, the preprocessing was performed as follows (lower trace of Fig. 3) [36–38]:

Duration: To be accepted as a correct respiratory cycle, the minimum duration of the cycle was set at 1500 ms. Shorter respiratory cycles were eliminated by the removal of a peak and a valley in such a way that the amplitude difference between the remaining successive peaks and valleys was maximal. It should be noted that the duration of respiratory cycles was defined from valley to valley.

Amplitude: To satisfy the detection criterion, the amplitude difference between a peak and a valley was supposed to be at least 15% of the previous and the following amplitude difference. Otherwise, the removal of a peak and a valley was performed in order to ensure that a maximal amplitude difference between the 2 optima was obtained.

2.4.2. Feature extraction from the R signal

Correa et al. compared directly recorded R signals (chest, abdomen and oronasal airflow) and EDR signals by measuring the mutual correlation coefficient and the spectral coherence. They observed that the EDR signal had the highest correlation with the chest (thoracic) R signal. They also showed that raw R signals did not have perfect correlation with themselves [39]. Considering these findings, we used the thoracic R signal recorded by inductive plethysmography bands and sampled at 10 Hz.

First, thoracic R signals were filtered by a 10th order Butterworth low pass filter with a 0.8 Hz cut-off frequency in order to remove high frequency noise and variations above the respiratory frequency. Then, a 6-point moving average filtering method was used to smooth out the thoracic R signal and remove additional peaks. Finally, since the signal was not calibrated in terms of absolute tidal volume, we normalized it for each subject and considered only relative differences. The thoracic R signal was normalized by first detecting the turning points and then calculating the difference between sequential peaks and troughs. The median peak-to-trough amplitude over the entire record was then determined and the signal was normalized by dividing into this value, so that the median peak-to-trough amplitude was unity [22]. The median was more robust to outliers and did not move (change) unless more than half of the signal contained noise, which in the plethysmogram can be extreme and create very large peak-to-trough values and can skew the mean.

Overall, features were extracted from 5-min segments of thoracic R signals in Stage 2, SWS, REM sleep and wakeful stages. First, peaks and valleys of thoracic R signals were detected. Then, the missing detections were corrected by the procedure explained in Section 2.4.1. Finally, the tidal volume (TV) and the respiration rate (RR) were calculated as the amplitude difference between a

Table 1
The number of 5-min segments of ECG and thoracic R signals for 30 subjects (REM and SWS stand for rapid eye movement and slow wave sleep, respectively).

Total segments	REM	SWS	Stage 2	Wake	
957	259	185	434	79	*The number of continuous 5-min segments*

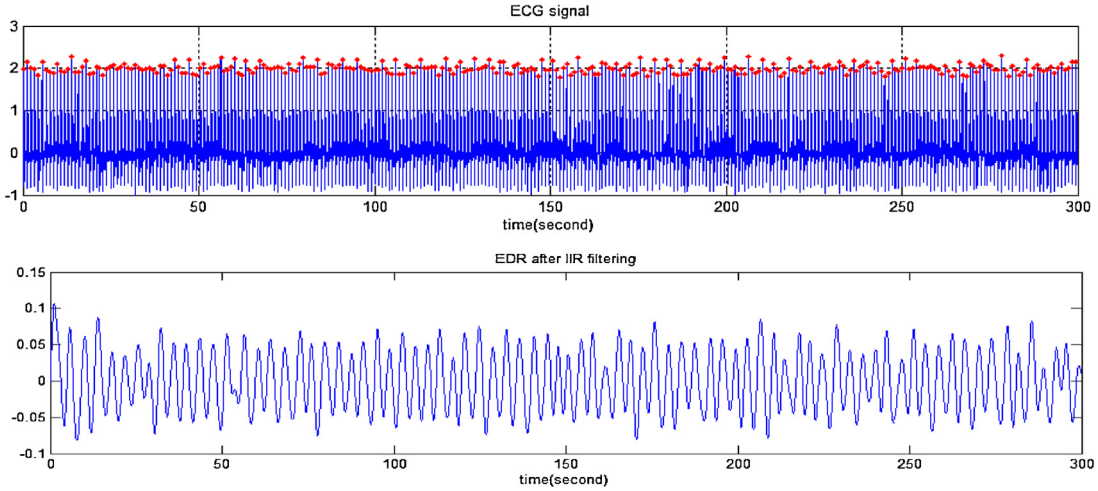

Fig. 2. A 5-min segment of ECG signal (upper trace). The RS-EDR signal derived from the ECG signal (lower trace).

successive peak and valley, and the number of breaths per minute, respectively. In this study, the mean, the standard deviation and the coefficient of variation (standard deviation 100/mean) of the TV and RR were computed as features.

Other features, such as Approximate Entropy (ApEn), Sample Entropy (SampEn), and Shannon Entropy (ShanEn) were extracted from thoracic R signals. Entropy is one of the most important parameters in biological signal processing, which extracts the information generation rate in these systems. For an N-point segment of a time series, with an embedding dimension m and a similarity parameter r, ApEn is given by the following [40]:

$$\mathrm{ApEn}(m, r, N) = \varphi^m(r) - \varphi^{m+1}(r) \tag{2}$$

$$\varphi^m(r) = [N-(m-1)\tau]^{-1} \sum_{i=1}^{N-(m-1)\tau} \ln C_i^m(r)$$

where

$$C_i^m(r) = \frac{B_i}{N-(m-1)\tau} \tag{3}$$

$$B_i = \text{number of } j \text{ such that } d|X_i, X_j| \le r \tag{4}$$

In the above equation (X_i, X_j) are m-dimensional pattern vectors, whose components are time delayed versions of the elements in the original time series with delay τ, a multiple of the sampling time, as follows:

$$X_i = (x_i, x_{i+\tau}, x_{i+2\tau}, \ldots, x_{i+(m-1)\tau}) \quad X_i \in R^m$$

$$X_j = (x_j, x_{j+\tau}, x_{j+2\tau}, \ldots, x_{j+(m-1)\tau}) \quad X_j \in R^m$$

And $d|X_i, X_j|$ is a measure of the distance between X_i and X_j. For large values of N the ApEn is given by:

$$\mathrm{ApEn}(m, r, N) = [N-m\tau]^{-1} \sum_{i=1}^{N-m\tau} -\ln\left(\frac{A_i}{B_i}\right) \tag{5}$$

where A_i is the number of X_i within tolerance r of X_j for the $(m+1)$ – dimensional pattern vector and B_i is the number of X_i with tolerance r of X_j in the m-dimensional pattern vector.

SampEn is another measure of complexity [40], which is very similar to ApEn. The main difference between these two measures is how self-counting is handled in their computation. In calculation of the ApEn, self-counting is included at each iteration to prevent computing the natural logarithm of zero. However, in SampEn calculation, the natural logarithm is computed once and self-counting

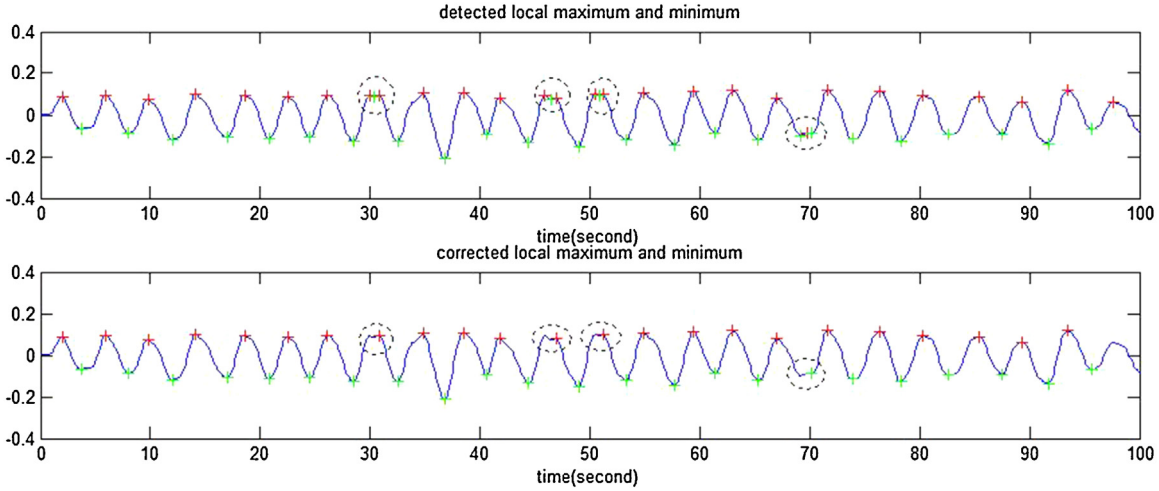

Fig. 3. An R signal after local optima detection (upper trace). False detections of local optima were removed in order to find the global optima (lower trace), and marked with small ellipses around them.

is excluded by requiring that $i \neq j$ in Eq. (5). SampEn is computed by modifying the ApEn formula given in Eq. (5) to:

$$\text{SampEn}(m, r, N) = -\ln\frac{A}{B} = -\ln\frac{\Sigma_{i=1}^{N-m\tau} A_i}{\Sigma_{i=1}^{N-m\tau} B_i} \quad (6)$$

The calculation of ApEn and SampEn requires a priori specification of some unknown parameters: m, the embedding dimension (ED) – r, a tolerance value and τ, time delay. In our study, we considered $m = 2$, $r = 0.2$ times the standard deviation of the data and $\tau = 11$ [12]. Time delay was determined as the lag at the point at which the autocorrelation function of the respiratory signal was near zero for the first time.

Finally, the peak frequency of thoracic R signals and the power value in peak frequency were extracted. The Power Spectral Densities (PSDs) of 5-min segments of R signals were estimated by using the nonparametric method and 1024-point FFT (Fast Fourier Transform). Totally, 11 features were extracted from the 5-min segments of thoracic R signals.

The procedure for the feature extraction from RS-EDR signals was the same as that for thoracic R signals.

2.5. Feature extraction from HRV signals

Generally, a 4-feature set was extracted from the HRV signals in 5-min segments as follows: (1) the time domain features included: the median, Inter-Quartile Range (IQR), Mean Absolute Difference (MAD), mean, standard deviation and range; (2) the nonlinear dynamic features included: Detrended Fluctuation Analysis (DFA)-based features (α_1 and α_2, the details of which were presented in [41]) and entropy measures (Shannon entropy, ApEn and SampEn); (3) the feature extraction by DWT, which included: the normalized values of energy in VLF (very low frequency), LF and HF bands (Wave 1:3) – the ratio of energies in LF/HF (Wave 4) – Shannon entropy in VLF, LF and HF bands (Wave 5:7) – the ratio of entropies in LF/HF(Wave 8); (4) feature extraction by the Empirical Mode Decomposition (EMD) method consisted of normalized values of energy in VLF, LF and HF bands computed by the Hilbert energy spectrum (EMD 1:3) – the ratio of energies in LF/HF (EMD 4) – harmonic parameters such as the central frequency, the deviation of central frequency and the energy in central frequency were extracted from the Hilbert amplitude spectrum (EMD 5–7) – ApEn (EMD 8:11) and SampEn (EMD 12:15) were calculated from the 4 most significant IMFs of the 5-min segments. To calculate ApEn and SampEn from HRV signals, the following values were selected: $m = 2$, $r = 0.2$ times the standard deviation of the data, and $\tau = 1$ [42]. Time-frequency methods are suitable for the extraction of cyclical variations. Detrended Fluctuation Analysis is mainly employed to analyze the scaling behavior and detect long-range correlations and irregularities of signals are measured by the calculation of entropy. As a result, it can be safely claimed that these methods are complementary techniques for HRV signal analysis.

2.6. Classification and optimization

SVMs are designed for binary-classification problems, assuming the data are linearly separable. Given the training data (x_i, y_i), $i = 1, \ldots, l$, $x_i \in R^n$, $y_i \in \{+1, -1\}$, where R^n is the input space, x_i is the sample vector and y_i is the class label of x_i, the separating hyperplane (ω, b) is a linear discriminating function that solves the optimization problem given as:

$$\min_{\omega,b} \langle \omega, \omega \rangle \quad (7)$$

Subject to $yi(\langle \omega, x_i \rangle + b) \geq 1$, $i = 1, 2, \ldots, l$where $\langle .,. \rangle$ indicates the inner product operation. The minimal distance between the samples and the separating hyperplane, i.e. the margin, is $1/||\omega||$.

In order to relax the margin constraints for the non-linearly separable data, the slack variables are introduced into the optimization problem:

$$\min_{\xi,\omega,b} \langle \omega, \omega \rangle + C\sum_{i=1}^{l} \xi_i \quad (8)$$

subject to $yi(\langle \omega, x_i \rangle + b) \geq 1 - \xi_i$, $i = 1, 2, \ldots, l$, $\xi_i \geq 0$

This leads to a soft margin SVM that is generally discussed and applied. The resulting classifier is called the 1-norm soft margin SVM, and C is the penalty parameter of error. The decision function of the classifier is:

$$\text{sign}(\Sigma_{x_i:SV} y_i \alpha_i \langle x_i, x \rangle + b) \quad (9)$$

In practice, since real data are often not linearly separable in the input space, the data can be mapped onto a high dimensional feature space, in which they are sparse and possibly more separable. When using a function φ: $X \rightarrow F$ to map the data onto a high dimensional feature space, the decision function of the classifier becomes:

$$\text{sign}(\Sigma_{x_i:SV} y_i \alpha_i \langle \varphi(x_i), \varphi(x) \rangle + b) \quad (10)$$

The mapping is often not explicitly given. Instead, a kernel function $(K(x_1, x_2) = \langle \varphi(x_1), \varphi(x_2) \rangle)$ is incorporated to simplify the computation of the inner product value of the transformed data in the feature space and the decision function becomes:

$$\text{sign}(\Sigma_{x_i:SV} y_i \alpha_i K(x_i, x) + b) \quad (11)$$

When the data are linearly inseparable, a non-linear kernel that maps the data onto the feature space non-linearly can handle the data better than the linear kernels. Gaspar et al. showed that the most common kernels in the literature: linear, polynomial and radial basis function (RBF), yield the best improvements [43]. The RBF kernel $(k(x, y) = e^{-\gamma||x-y||^d})$ is a reasonable first choice of kernel function [44]. When using the RBF kernel, the parameters d, γ should be set properly. Generally d is set to be 2. Thus the kernel value is related to the Euclidean distance between the two samples, and γ is related to the kernel width. To apply SVM to multi-class classification problems, the problem can be divided into sub-problems, which are binary classification problems. The often suggested implementations for SVM multi-class classification are the one-against-rest and the one-against-one method. Lin et al. [45] showed that the one-against-one method performs the best and can be trained faster than the SVM models in the one-against-rest method.

In our study, the SVM separating hyperplane was calculated by solving the quadratic optimization problem. The RBF kernel and the one-against-one method were used for SVM multi-class classification. The dataset was divided into the training (80%) and the test set (20%). Then, the 10-fold cross-validation was used to find optimal parameters (C, γ). In the 10-fold cross-validation, first the training set was divided into 10 subsets of equal size. Sequentially one subset of 10 training subsets was tested using the classifier trained on the remaining 9 subsets. Thus, each instance of the whole training set was predicted once so the cross-validation accuracy was the percentage of data, which were correctly classified. The cross-validation procedure can prevent the over-fitting problem. A "grid-search" on C and γ ($C = 2^{-5}\ 2^{-3}, \ldots, 2^{15}$, $\gamma = 2^{-15}; 2^{-13}, \ldots, 2^{3}$) was used to find optimal pairings [44]. Various pairs of (C, γ) values were tried and the one with the best cross-validation accuracy was picked. Finally, the SVM was trained by optimal parameters using all the training data (80% feature vectors) and was then tested on the test data (20% feature vectors).

Feature reduction has an important role in classification. It reduces computational load and yields good performance for the

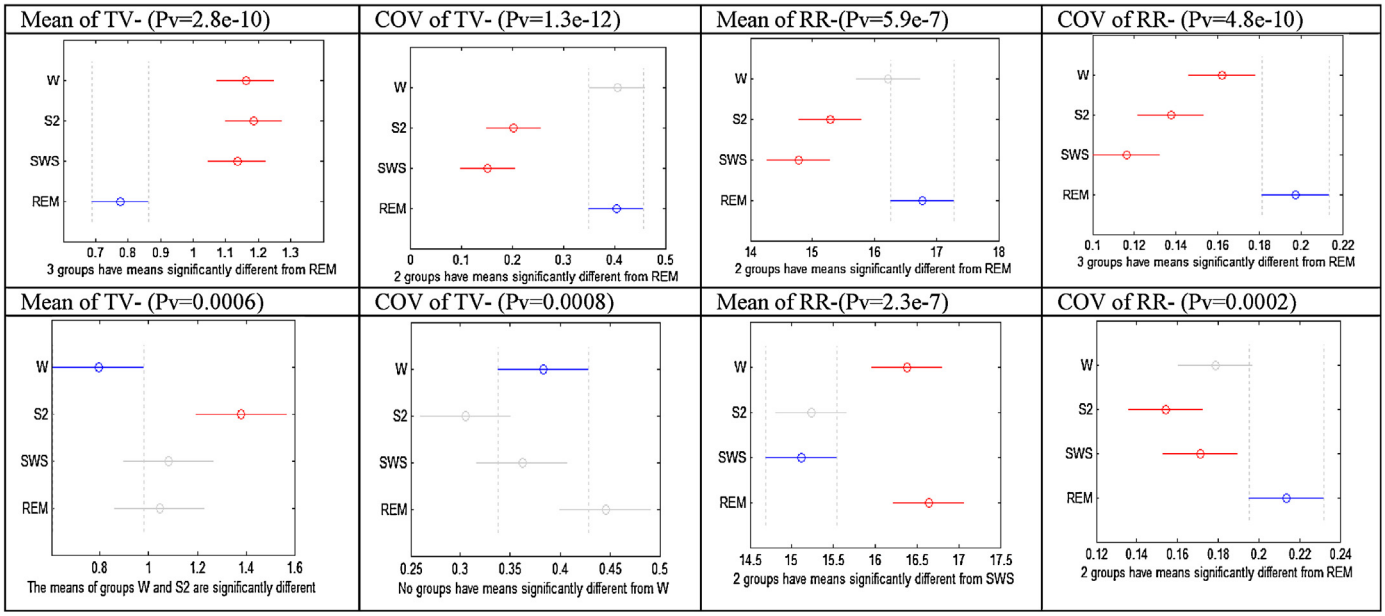

Fig. 4. The results of the ANOVA test and a post hoc Bonferroni correction for 4 extracted features of thoracic R signals (upper parts) and RS-EDR (lower parts). TV, tidal volume; RR, respiration rate; COV, coefficient of variation; Pv, probability value of ANOVA.

classifier. Generally, there are three feature selection methods: embedded, filter and wrapper. In embedded-based methods, feature selection is a part of the classifier training process and these methods exploit the knowledge of the specific structure of the learning algorithm. In filter-based methods, a subset of features is selected based on evaluation criterion which is independent of the learning algorithm, and in the wrapper methods, feature selection is based on performance of the classifier [46]. Here, we used the SVM Recursive Feature Elimination (RFE) method, which is a wrapper-based method. The SVM-RFE was developed by Guyon et al., and has been used in gene selection for cancer classification [47]. The description of this method is given below.

In the SVM-RFE method, the effect of removing a feature on an objective function is used as a ranking criterion. Guyon, et al. used the weight magnitude of support vectors as ranking criterion [47]. Adnane et al. used the ranking criterion Total Error Rate (TER) [19]. Here, we use the ranking criterion Correct Classification Ratio (CCR). The steps of the SVM-RFE feature selection algorithm are as follows:

(a) Remove one feature out of the number of features (NF) with an initial value of NF = 45 and compute the ranking criterion (calculated on test data). This operation is repeated for every feature removed.
(b) Compare the ranking criterion values obtained for each subset of (NF − 1) features and sort out the feature with the highest ranking criterion, which is removed completely. A new subset of (NF − 1) features is then treated in step (a). In the case of CCR used as ranking criterion (as we have done here), the subset of (NF − 1) features with the higher CCR contains the best features, and the feature that was removed in step (a) from the set (or subset) of NF features is considered as the worst one.

3. Results

We used 3 sets of features extracted from: HRV, combined HRV+ RS-EDR, and combined HRV+ thoracic R signals for automatic sleep staging. We evaluated the importance of the extracted features from HRV signals by statistical methods and published before [41]. In this study, we extracted 11 features from either RS-EDR or thoracic R signals. The ANOVA test followed by a post hoc Bonferroni correction was used for individual feature assessment. The results of the ANOVA test for the mean and coefficient of variation (COV) of thoracic R signals and RS-EDR are presented in Fig. 4.

We observed that the variability of RR is the same between thoracic R signals and RS-EDR. This means that the mean value of the RR decreases from wakefulness to NREM sleep and increases in REM sleep. Also, the standard deviation and COV of RR are high in REM sleep. The mean value of the TV derived from the thoracic R signals has the minimum value in REM sleep, which means that the respiration depth decreases in REM sleep. All of these findings are consistent with respiration physiology in sleep stages [5] but as for the RS-EDR signal, there was an exception – the mean value of the TV was at a minimum in wakeful stage.

First, automatic sleep staging was performed by using 3 sets of features (HRV alone, combined HRV+ RS-EDR signals, and combined HRV+ thoracic R signals). We used the SVM classifier, and optimal parameters for this classifier were found by the procedure described in Section 2.6. The results of classification by using the optimal parameters of the SVM classifier are presented in Tables 2–4.

Table 4 shows the best results obtained by using a combination of HRV and thoracic R features. Evidently, adding the RS-EDR features to the HRV features improved the classification results (the results shown in Table 3 are better than those shown in Table 2).

Table 2
Automatic sleep staging by using HRV signals alone. Results are reported for optimum $C = 128$ and $\sigma = 32$ ($\gamma = 0.00097$). In addition, CCR1 (on training data) = 0.68 and CCR2 (on test data) = 0.57.

Sleep stages	Wake	Stage 2	SWS	REM	Total on test data	Total on training data
Accuracy	80.96	65.18	67.99	69.17	76.56	81.50
Specificity	86.93	80.95	81.93	86.42	84.37	87.66
Sensitivity	75	49.42	54.05	51.92	53.12	63.00

Table 3
Automatic sleep staging by using HRV+ RS-EDR signals. Results are reported for optimum $C = 2048$ and $\sigma = 64$ ($\gamma = 0.00024$). In addition, CCR1 (on training data) = 0.76 and CCR2 (on test data) = 0.65.

Sleep stages	Wake	Stage 2	SWS	REM	Total on test data	Total on training data
Accuracy	92.00	70.83	80.72	80.72	81.25	85.62
Specificity	93.75	78.09	87.09	87.14	87.50	90.41
Sensitivity	81.25	62.06	54.05	63.46	62.50	71.24

Table 4
Automatic sleep staging by using HRV+ thoracic R signals. Results are reported for optimum $C = 8$ and $\sigma = 8$ ($\gamma = 0.0156$). In addition, CCR1 (on training data) = 0.91 and CCR2 (on test data) = 0.76.

Sleep stages	Wake	Stage 2	SWS	REM	Total on test data	Total on training data
Accuracy	94.27	78.64	86.97	89.06	**87.23**	94.77
Specificity	95.45	82.85	90.32	94.28	**91.49**	96.51
Sensitivity	81.25	73.56	72.97	75.00	**74.47**	89.54

The best obtained results of automatic sleep staging by cardiorespiratory signals on test samples are bold.

Table 5
The best 35 features selected by using the SVM-RFE ranking method (features were extracted from HRV and RS-EDR signals). The 9 features extracted from RS-EDR signals are highlighted.

Rank	Feature description
1	'Standard deviation, HRV'
2	'Mean, HRV'
3	'DFA 2, HRV'
4	'Wave 3, HRV'
5	'Respiration rate-SD, RS-EDR'
6	'Tidal volume-mean, RS-EDR'
7	'Sample entropy, RS-EDR'
8	'EMD 11, HRV'
9	'Wave 1, HRV'
10	'Wave 8, HRV'
11	'EMD 9, HRV'
12	'Tidal volume-standard deviation, RS-EDR'
13	'EMD 13, HRV'
14	'DFA 1, HRV'
15	'Wave 7, HRV'
16	'EMD 10, HRV'
17	'Wave 5, HRV'
18	'EMD 14, HRV'
19	'IQR, HRV'
20	'Respiration rate-mean, RS-EDR'
21	'Amplitude in peak frequency, RS-EDR'
22	'EMD 6, HRV'
23	'EMD 8, HRV'
24	'Shannon entropy, RS-EDR'
25	'EMD 5, HRV'
26	'EMD 1, HRV'
27	'ApEn, RS-EDR'
28	'Sample entropy, HRV'
29	'Wave 4, HRV'
30	'ApEn, HRV'
31	'EMD 7, HRV'
32	'EMD 3, HRV'
33	'EMD 12, HRV'
34	'Wave 2, HRV'
35	'Respiration rate – coefficient of variation, RS-EDR'

Building upon these outcomes, we attempted to improve the classification results by refining our feature selection. We used the SVM-RFE method (the details of which are described in Section 2.6) for this purpose.

Then, the SVM-RFE ranking method was applied to 45 features extracted from HRV and RS-EDR signals. The best result was obtained when using 35 features which overall are the best feature set. The results of ranking are summarized in Table 5 (ranked from the best to the worst outcomes). The numbers of 9 features from the best feature set are from RS-EDR signal, which are highlighted for better visualization.

Fig. 5 shows the CCR versus the number of best features in an increasing fashion. It is clear that the best results were obtained

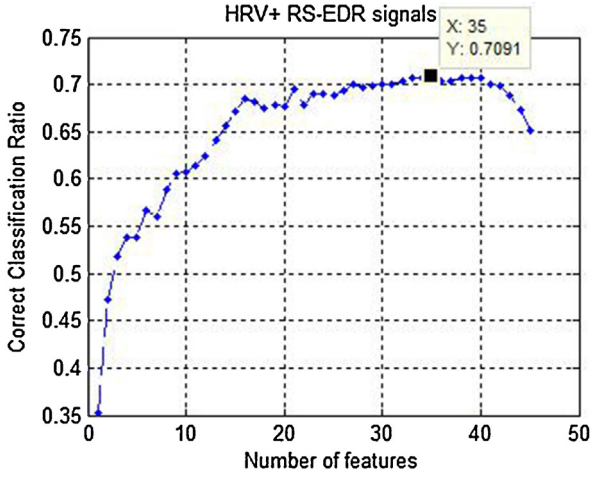

Fig. 5. From right to left, each point in this diagram is the best possible subset (which maximize CCR) of m features (m is changed from 45 to 1). When a feature is removed, all possible '$n-1$' combinations of 'n' remaining features is checked, it means that removed the features in previous steps cannot be tested in next combinations. The best result was obtained by the combination of 35 features.

for the set of 35 features. The classification results for the best 35 features derived from the combined HRV and RS-EDR signals are presented in Table 6.

The details of the SVM-RFE method are presented in final paragraph of Section 2.6. The main point of this method is that by an initial number of 'n' features, the SVM-RFE checks all possible '$n-1$' combinations from the number of 'n' features. Then selects the best combination, which maximizes CCR. In the next step, this process is repeated by '$n-1$' features. It continues till that one feature remains. In this study, first the best combination of 44 features was found. In next step, the best combination of 43 features was found... till that in the two final steps, the best combination of 2 features was determined, and finally the best feature was found. Therefore, each point in diagram of Fig. 5 is the best possible subset (which maximizes CCR) of m features (m is changed from 45 to 1). Generally, when a feature is removed, all possible '$n-1$' combinations of 'n' remaining features are checked. This means that the removed features in the previous steps cannot be tested in next combinations. By evaluating the diagram from the end point to the first point, CCR is increased gradually by reducing features from 45 to 35 features. As more features are removed, the value of CCR decreases. The best result was obtained by the combination of 35 features. Please note that at some points such as $m = 7$ in Fig. 5, there are local minima or maxima. It should also be noted that the features that are top ranked (eliminated last) are not necessarily the

Table 6
Automatic sleep staging by using HRV+ RS-EDR signals. Results are reported for optimum $C = 2048$ and $\sigma = 64$ ($\gamma = 0.00024$) when using the 35 best features. In addition, CCR1 (on training data) = 0.73 and CCR2 (on test data) = 0.7.

Sleep stages	Wake	Stage 2	SWS	REM	Total on test data	Total on training data
Accuracy	92.70	75.00	83.85	83.85	83.85	84.11
Specificity	93.18	82.85	89.67	88.57	89.23	89.41
Sensitivity	87.50	65.51	59.45	71.15	67.7	68.23

Table 7
The best 27 features selected by using the SVM-RFE ranking method (features were extracted from HRV and thoracic R signals.) The 9 features extracted from thoracic R signals are highlighted.

Rank	Feature description
1	'Respiration rate – coefficient of variation, R signal'
2	'EMD 3, HRV'
3	'Tidal volume – standard deviation, R signal'
4	'Respiration rate – mean, R signal'
5	'Standard deviation, HRV'
6	'Power in peak frequency, R signal'
7	'Wave 8, HRV'
8	'Amplitude in peak frequency, R signal'
9	'Tidal volume – mean, R signal'
10	'Sample entropy, HRV'
11	'Sample entropy, R signal'
12	'Respiration rate – standard deviation, R signal'
13	'DFA 1, HRV'
14	'Median, HRV'
15	'DFA 2, HRV'
16	'Wave1, HRV'
17	'EMD 6, HRV'
18	'Wave 3, HRV'
19	'Mean, HRV'
20	'EMD14, HRV'
21	'Shannon entropy, HRV'
22	'EMD12, HRV'
23	'EMD 2, HRV'
24	'ApEn, R signal'
25	'EMD 5, HRV'
26	'EMD10, HRV'
27	'IQR, HRV'

ones that are individually most relevant. Only taken together the features of a subset are optimal in some sense [47].

Similarly, the SVM-RFE ranking method was applied to 45 features extracted from HRV and thoracic R signals. The best result was obtained when using 27 features which overall are the best feature set. The results of ranking are summarized in Table 7 (ranked from the best to the worst outcomes). The numbers of 9 features from the best feature set are from thoracic R signals, which are highlighted for better visualization.

Fig. 6 shows the CCR versus the number of best features in an increasing fashion. It is clear that the best results are obtained for the set of 27 features. The classification results for the best 27 features derived from the combined HRV and thoracic R signals are presented in Table 8.

Finally, two simple linear and nonlinear classifiers were used in automatic sleep staging based on the combination of HRV and thoracic R signals. These were Linear Discriminant (LD) and Nearest Neighbor (NN) Classifiers, respectively. The classification results are presented in Table 9. The low performances obtained by these

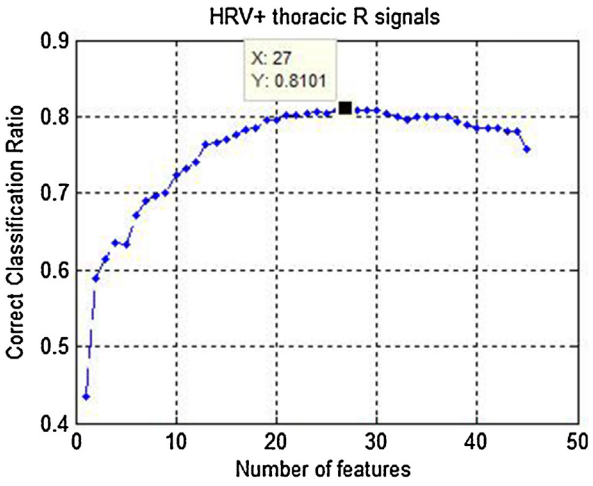

Fig. 6. From right to left, each point in this diagram is the best possible subset (which maximizes CCR) of m features (m is changed from 45 to 1). When a feature is removed, all possible '$n-1$' combinations of 'n' remaining features is checked, it means that the removed features in previous steps cannot be tested in next combinations. The best result was obtained by the combination of 27 features.

Table 9
Automatic sleep staging by using the combination of HRV and thoracic R signals. Results were presented on similar test samples, CCR (SVM) = 0.81, CCR (LD) = 0.58 and CCR (NN) = 0.53.

Classifiers	SVM classifier	LD Classifier	NN Classifier
Accuracy	89.32	82	81.5
Specificity	92.88	88	87.6
Sensitivity	78.64	64	63

simple classifiers clearly demonstrate the significance of the results achieved by the SVM classifier.

In our investigation, we used the CCR as a performance measure for classification results. We calculated CCR from the confusion matrix. After performing classification, we obtained a confusion matrix as shown in Table 10. We normalized each row of the confusion matrix by the number of examples in that class. The CCR is defined as the average of diagonal elements of the confusion matrix after normalization.

$$CR = \frac{(N_{11}/N_1) + (N_{22}/N_2) + (N_{33}/N_3) + (N_{44}/N_4)}{4} \quad (12)$$

4. Discussion

In this investigation, we compared the automatic sleep staging by using ECG-derived signals and the combination of HRV and R signals. The importance of HRV signals in automatic sleep staging

Table 8
Automatic sleep staging by using HRV+ thoracic R signals. Results are reported for optimum $C = 8$ and $\sigma = 8$ ($\gamma = 0.0156$) when using the 27 best features. In addition, CCR1 (on training data) = 0.83 and CCR2 (on test data) = 0.81.

Sleep stages	Wake	Stage 2	SWS	REM	Total on test data	Total on training data
Accuracy	95.83	81.25	89.06	91.14	89.32	90.58
Specificity	96.59	87.61	91.61	93.57	92.88	93.72
Sensitivity	87.50	73.56	78.37	84.61	78.64	81.17

Table 10
The confusion matrix or output of classification.

		Decided class by classifier				Total
		Wake	Stage 2	Slow wave sleep	REM	
Real class by gold standard	Wake	N_{11}	N_{12}	N_{13}	N_{14}	N_1
	Stage 2	N_{21}	N_{22}	N_{23}	N_{24}	N_2
	Slow wave sleep	N_{31}	N_{32}	N_{33}	N_{34}	N_3
	REM	N_{41}	N_{42}	N_{43}	N_{44}	N_4

was investigated in depth in our previous research [41]. We concluded that the extracted features from 5-min segments were more informative than those derived from 0.5-min epochs. Therefore, in this paper features were extracted from 5-min segments of cardiorespiratory (HRV, RS-EDR and reference R) signals, which is in accordance with the recommended record length for the analysis of these signals. The SVM-RFE wrapper feature selection method was used in order to improve the classification results based on the combination of HRV and RS-EDR features as well as the combination of HRV and thoracic R signals features. Finally, the SVM was used for classification. It was discovered that in classifying wake, Stage 2, slow wave sleep, and REM the best results were obtained when 27 features were extracted from HRV and thoracic R signals with a mean classification *accuracy of 89.32% and a CCR of 0.81* on test samples were used. It was also revealed that classification results by ECG-derived signals (HRV and RS-EDR signals) had a mean *accuracy of 83.85% and a CCR of 0.7*.

Our findings here contribute to the advancement of the field of automatic sleep staging using cardiorespiratory signals. The relevant literature shows that the best classification accuracies using optimum EEG-based subject-independent systems are in the range of 80–85%, when averaged over both normal and pathological populations [48–51]. Moreover, a number of other studies, in which cardiorespiratory signals were used, have been able to separate sleep and wake or REM and NREM stages. For example, Mendez et al. reported an accuracy of 79% for separating NREM and REM stages by using HRV signals. They used time-variant autoregressive models as feature extractor and a hidden Markov model as stage classifier [18]. Andane el al. also used HRV signals for sleep analysis. They extracted features by spectral, time domain and DFA methods, and could separate sleep and wakeful stages attaining 79.9% accuracy when using the SVM classifier [19]. In a very recent study, spectral parameters of the heart rate were used for the classification of wake, light sleep, deep sleep, and REM sleep. The reliability of the algorithm was tested with 18 ECGs and 57.7% of all 30 s epochs were correctly assigned by the algorithm [20]. Redmond et al. obtained an accuracy of 76.1% utilizing 27 features derived from RR intervals, EDR, and respiratory effort, to identify wake, NREM, and REM in a subject dependent manner. Their reported accuracy decreased to 67% when a subject independent system was designed [21]. They also obtained an accuracy of 76.1% for a 3-class (Wakefulness, NREM Sleep and REM) system using cardiorespiratory signals [22].

Most of the recent investigations on automatic sleep staging by cardiorespiratory signals did not evaluate all sleep stages separately. Some of the studies in literature are subject independent [22], and some others are subject specific [18,20,21]. In comparison to other investigations, it is noteworthy that our research produced improved discriminative capacity in distinguishing between Wakefulness, Sleep Stage2, slow wave sleep and REM. Moreover, this research was performed in a subject specific manner.

It has been shown by a number of researchers that automatic sleep staging based on feature vectors extracted from *0.5 min epochs* of cardiorespiratory signals, have not been so successful. Here we showed that feature extraction from *5-min long epochs* of HRV and reference R signals produced high discriminative capacity in classifying wake, stage2, SWS and REM sleep. We also showed that sleep staging by ECG-derived signals alone produced acceptable results. To summarize, in this investigation we used a combination of time domain, time-frequency and nonlinear methods for feature extraction. Our feature selection procedure demonstrated that a combination of time domain, time-frequency, and fractal features, in addition to entropies were necessary to reach the optimal result in performing automatic sleep staging by using cardiorespiratory signals. It was observed that the selection of optimal features not only improved the classification results but also increased the generalizability of the final outcomes as the results obtained from training and testing data sets became almost the same. To reiterate, the optimal feature set based on using the combination of HRV and R signals for sleep staging comprised of 27 features, 33% of which were derived from R signal. The 41% of the optimal feature set was extracted from HRV signals by nonlinear system analysis methods (EMD, DFA and measuring entropy). This high percentage shows the importance of using nonlinear methods in HRV signal analysis. The remaining 26% of the optimal feature set was extracted from HRV signals by time domain and wavelet transform methods.

A comparison between Tables 5 and 7 shows that the extracted features from reference R signals have higher ranking than those derived from RS-EDR features. Therefore, by improving the EDR extraction method, the results of sleep staging by ECG-based signals (the combination of HRV and EDR) alone will be comparable to those generated by the cardiorespiratory signals (the combination of HRV and reference R signal).

As with any algorithm used in automatic sleep staging, our method also has some limitations. Firstly, sleep scoring is routinely performed on 0.5-min segments of PSG data, which is a standard record length used for manual sleep staging in sleep labs [1,2]. However, our feature vectors had to be extracted from 5-min segments, which is the recommended record length for computer-based analysis of cardiorespiratory signals [52]. This limitation is somehow unavoidable as HRV and R signals are slow dynamic signals and it is reasonable to expect that more information could be obtained from 5-min segments than 0.5-min segments. Moreover, as no *continuous* 5-min epochs were available for sleep Stage 1 in the database that we used, Stage1 was not considered for further analysis.

Secondly, we used a subject-dependent classifier. The problem is that heart rate variations during wake and sleep stages are influenced differently by the ANS and this in turn changes from subject to subject. Although the idea of a subject-independent classifier for sleep staging is motivating, it is not feasible, at least when cardiorespiratory signals are used for sleep–wake stage classification. This is confirmed by the low performance results obtained by Redmond and Heneghan [21] who developed a subject-independent system, even though they utilized 36 data records (all data records except the tested data records) as training data.

5. Conclusion

In this study, an automatic sleep staging algorithm was developed and tested using two different sets of signals: (1) ECG-derived (HRV and RS-EDR) and (2) cardiorespiratory (the combination of HRV and R – thoracic excursion). First, a set of various features

was extracted from long epochs of ECG-derived and R signals (each of 5-min length) using a combination of advanced linear and nonlinear methods. Then, the SVM-RFE wrapper feature selection method was employed to select the most effective features in order to improve the classification results. Finally, the sleep stages were classified by means of the SVM algorithm.

Our results demonstrated the effectiveness of combining time domain, time-frequency, and fractal features, in addition to the entropies in automatic sleep staging by means of cardiorespiratory signals. It was also shown that by selecting an optimal set of features it is not only possible to improve the classification results but also to increase the generalization property of the classifier. The best results in classifying the four classes of Wake, Stage 2, slow wave sleep, and REM were obtained when only 27 features of those extracted from HRV and thoracic R signals were used. It can be concluded that sleep staging by only ECG-derived signals can produce acceptable results.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgment

The authors would like to thank the SHHS for providing the polysomnographic data.

References

[1] A. Rechtschaffen, A. Kales, A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects, Public Health Service, S. Government Printing Office, Washington, DC, 1968.
[2] C. Iber, S. Ancoli-Israel, A. Chesson, S.F. Quan, The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications for the American Academy of Sleep Medicine, American Academy of Sleep Medicine, Westchester, 2007.
[3] H. Nazeran, Y. Pamula, K. Behbehani, Heart Rate Variability (HRV): Sleep Disorder Breathing, Wiley Encyclopedia in Biomedical Engineering, Wiley, New York, 2006, pp. 1–17.
[4] G. Mancia, Autonomic modulation of the cardiovascular system during sleep, N. Engl. J. Med. 328 (1993) 347–349.
[5] W. Moorcroft, P. Belcher, Understanding Sleep and Dreaming; The Body During Sleep, Part II, Springer, New York, 2005, pp. 97–119.
[6] D. Vigo, J. Dominguez, S. Guinjoan, M. Scaramal, E. Ruffa, J. Solerno, L.N. Siri, D. Cardinali, Nonlinear analysis of heart rate variability within independent frequency components during the sleep–wake cycle, Auton. Neurosci.: Basic Clin. 154 (2010) 84–88.
[7] P. Busek, J. Vankova, J. Opavsky, J. Salinger, S. Nevsimalova, Spectral analysis of heart rate variability in sleep, Physiol. Res. 54 (2005) 369–376.
[8] T. Penzel, J.W. Kantelhardt, L.C. Chang, K. Voigt, C. Vogelmeier, Dynamics of heart rate and sleep stages in normals and patients with sleep apnea, Neuropsychopharmacology 28 (2003) 48–53.
[9] T. Penzel, J.W. Kantelhardt, L. Grote, J.H. Peter, A. Bunde, Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea, IEEE Trans. Biomed. Eng. 50 (2003) 1143–1151.
[10] M.A. Carskadon, K. Elarvey, W.C. Dement, et al., Respiration during sleep in children, West. J. Med. 128 (1978) 477–481.
[11] M. Miyata, N. Burioka, H. Suyama, T. Sako, T. Nomura, T. Takeshima, S. Higami, E. Shimizu, Non-linear behaviour of respiratory movement in obstructive sleep apnœa syndrome, Clin. Physiol. Funct. Imaging 22 (2002) 320–327.
[12] N. Burioka, G. Cornelissen, F. Halberg, D.T. Kaplan, H. Suyama, T. Sako, E. Shimizu, Approximate entropy of human respiratory movement during eye-closed waking and different sleep stages, Chest 123 (2003) 80–86.
[13] P.J. Fadel, S.M. Barman, S.W. Phillips, G.L. Gebber, Fractal fluctuations in human respiration, J. Appl. Physiol. 97 (2004) 2056–2064.
[14] P.D. Larsen, D.E. Elder, Y.C. Tzeng, A.J. Campbell, D.C. Galletly, Fractal characteristics of breath to breath timing in sleeping infants, Respir. Physiol. Neurobiol. 139 (2004) 263–270.
[15] M. Wysocki, M.-N. Fiamma, C. Straus, C.-S. Poon, T. Similowski, Chaotic dynamics of resting ventilatory flow in humans assessed through noise titration, Respir. Physiol. Neurobiol. 153 (2006) 54–65.
[16] R. Bartsch, J. Kantelhardt, T. Penzel, S. Havlin, Experimental evidence for phase synchronization transitions in the human cardiorespiratory system, Phys. Rev. Lett. 98 (2007) 54–102.
[17] R. Bartsch, A. Schumann, J. Kantelhardt, T. Penzel, P. Ivanov, phase transitions in physiologic coupling, Proc. Natl. Acad. Sci. 109 (2012) 10181–10186.
[18] M. Mendez, M. Matteucci, V. Castronovo, L. Ferini-Strambi, S. Cerutti, A. Bianchi, Sleep staging from heart rate variability: time-varying spectral features and Hidden Markov Models, Int. J. Biomed. Eng. Technol. 3 (2010) 246–263.
[19] M. Adnane, Z. Jiang, Z. Yan, Sleep–wake stages classification and sleep efficiency estimation using single-lead electrocardiogram, Exp. Syst. Appl. 39 (2012) 1401–1413.
[20] K. Kesper, S. Canisius, T. Penzel, T. Ploch, W. Cassel, ECG signal analysis for the assessment of sleep-disordered breathing and sleep pattern, Med. Biol. Eng. Comput. 50 (2012) 135–144.
[21] S.J. Redmond, C. Heneghan, Cardiorespiratory-based sleep staging in subject with obstructive sleep apnea, IEEE Trans. Biomed. Eng. 53 (2006) 1–12.
[22] S.J. Redmond, P. Chazal, C. Brien, S. Ryan, W. McNicholas, C. Heneghan, Sleep staging using cardiorespiratory signals, Somnologie 11 (2007) 245–256.
[23] W. Karlen, C. Mattiussi, D. Floreano, Sleep and wake classification with ECG and respiratory effort signals, IEEE Trans. Biomed. Circuits Syst. 3 (2009) 71–78.
[24] S.F. Quan, B.V. Howard, C. Iber, et al., The sleep heart health study: design, rationale, and methods, Sleep 20 (1997) 1077–1085, available from: https://sleepepi.partners.org/hybrid/
[25] S. Javaheri, T. Parke, J. Liming, et al., Sleep apnea in 81 ambulatory male patients with stable heart failure. Types and their prevalences, consequences, and presentations, Circulation 97 (1998) 2154–2159.
[26] R. Mehra, E. Benjamin, E. Shahar, et al., Association of nocturnal arrhythmias with sleep-disordered breathing: the sleep heart health study, Am. J. Respir. Crit. Care Med. 173 (2006) 910–916.
[27] A. Baranchuk, C.S. Simpson, D.P. Redfearn, K. Michael, M. Fitzpatrick, Understanding the association between sleep apnea and cardiac arrhythmias, Rev. Electrofisiol. Arritm. 1 (2008) 5–6.
[28] A.S. Hersi, Obstructive sleep apnea and cardiac arrhythmias, Ann. Thorac. Med. 5 (2010) 10–17.
[29] K. Monahan, A. Storfer-Isser, R. Mehra, E. Shahar, M. Mittleman, J. Rottman, N. Punjabi, M. Sanders, S. Quan, H. Resnick, S. Redline, Triggering of nocturnal arrhythmias by sleep-disordered breathing events, J. Am. Coll. Cardiol. 54 (2009) 1797–1804.
[30] A. Baranchuk, Sleep apnea, cardiac arrhythmias, and conduction disorders, J. Electrocardiol. 45 (2012) 508–512.
[31] D. Benitez, P.A. Gaydecki, A. Zaidi, A.P. Fitzpatrick, The use of the Hilbert transform in ECG signal analysis, Comput. Biol. Med. 31 (2001) 399–406.
[32] S. Chatlapalli, H. Nazeran, V. Melarkod, R. Krishnam, E. Estrada, Y. Pamula, S. Cabrera, Accurate derivation of heart rate variability signal for detection of sleep disordered breathing in children, in: Proceedings of the 26th Annual International Conference of the IEEE EMB Sin, San Francisco, USA, 1–5 September, 2004, pp. 538–541.
[33] M.P. Tarvainen, P.O. Ranta-aho, P.A. Karjalainen, An advanced detrending method with application to HRV analysis, IEEE Trans. Biomed. Eng. 49 (2002) 172–175.
[34] P. Chazal, T. Penzel, C. Heneghan, Automated detection of obstructive sleep apnœa at different time scales using the electrocardiogram, Physiol. Meas. 25 (2004) 967–983.
[35] C.L. Mason, L. Tarassenko, Quantitative assessment of respiratory derivation algorithms, in: IEEE EMBS Conference, Istanbul, Turkey, 25–28 October, 2001, pp. 1998–2001.
[36] D. Widjaja, J. Taelman, S. Vandeput, M. Braeken, R. Otte, B. Bergh, S. Huffe, ECG-derived respiration: comparison and new measures for respiratory variability, in: Computing in Cardiology, 2010, September, pp. 149–152.
[37] S. Babaeizadeh, S.H. Zhou, S.D. Pittman, D.P. White, Electrocardiogram-derived respiration in screening of sleep-disordered breathing, J. Electrocardiogr. 44 (2011) 700–706.
[38] B. Mazzanti, C. Lamberti, J. de Bie, Validation of an ECG-derived respiration monitoring method, Comput. Cardiol. 30 (2003) 613–616.
[39] L.S. Correa, E. Laciar, A. Torres, R. Jane, Performance evaluation of three methods for respiratory signal estimation from the electrocardiogram, in: 30th Annual International IEEE EMBS Conference, Vancouver, British Columbia, Canada, 20–24 August, 2008.
[40] J.S. Richman, J.R. Moorman, Entropy and sample entropy physiological time-series analysis using approximate? Am. J. Physiol. Heart Circ. Physiol. 278 (2000) 2039–2049.
[41] F. Ebrahimi, S.K. Setarehdan, J.A. Moyeda, H. Nazeran, Automatic sleep staging using empirical mode decomposition, discrete wavelet transform, time-domain, and nonlinear dynamics features of heart rate variability signals, Comput. Methods Progr. Biomed. 112 (2013) 47–57.
[42] K. Chon, C.G. Scully, S. LU, Approximate entropy for all signals: is the recommended threshold value r appropriate? IEEE Eng. Med. Biol. Mag. 28 (2009) 18–23.
[43] P. Gaspar, J. Carbonell, J.L. Oliveira, On the parameter optimization of support vector machines for binary classification, J. Integr. Bioinform. 9 (2012) 201–212.
[44] C.-W. Hsu, C.-C. Chang, C.-J. Lin, A Practical Guide to Support Vector Classification (Online), 2003, available from: http://www.csie.ntu.edu.tw/cjlin/libsvm (Last updated: April 15, 2010).
[45] C.-W. Hsu, C.-J. Lin, A comparison of methods for multiclass support vector machines, IEEE Trans. Neural Netw. 13 (2002) 415–425.

[46] K. Shen, C. Ong, X. Li, E. Wilder-Smith, Feature selection via sensitivity analysis of SVM probabilistic outputs, Mach. Learn. 70 (2008) 1–20.
[47] I. Guyon, J. Weston, S. Barnhill, V. Vapnick, Gene selection for cancer classification using support vector machines, Mach. Learn. 46 (2002) 389–422.
[48] A. Flexer, G. Gruber, G. Dorffner, Continuous unsupervised sleep staging based on single EEG signal Lecture Notes in Computer Science, Artificial Neural Network, vol. 2415, Springer Verlag, Berlin, Germany, 2002, pp. 1013–1018.
[49] T. Katayama, E. Suzuki, M. Saito, Staging of awake and sleep based on feature map, Syst. Comput. Jpn. 26 (1995) 98–107.
[50] C.L. Albertario, S.M. Zendell, G. Hertz, M.M. Maberino, S.H. Feinsilver, Comparison of a frequency-based analysis of electroencephalograms (*Z*-ratio) and visual scoring on the multiple sleep latency test, Sleep 18 (1995) 836–843.
[51] N. Schaltenbrand, R. Lengelle, M. Toussaint, R. Luthringer, G. Carelli, A. Jacqmin, E. Lainey, A. Muzet, J.P. Macher, Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients, Sleep 19 (1996) 26–35.
[52] Task Force of the European Society of Cardiology and North American Society of Pacing and Electrophysiology, Heart rate variability, standards of measurement, physiological interpretation and clinical use, Eur. Heart J. 17 (1996) 354–381.